

Grundlagen der Künstlichen Intelligenz

3. Einführung: Rationale Agenten

Malte Helmert

Universität Basel

2. März 2015

Grundlagen der Künstlichen Intelligenz

2. März 2015 — 3. Einführung: Rationale Agenten

3.1 Agenten

3.2 Rationalität

3.3 Zusammenfassung

Einführung: Überblick

Kapitelüberblick Einführung:

- ▶ 1. Was ist Künstliche Intelligenz?
- ▶ 2. KI früher und heute
- ▶ **3. Rationale Agenten**
- ▶ 4. Umgebungen und Problemlösungsverfahren

3.1 Agenten

Heterogene Einsatzgebiete

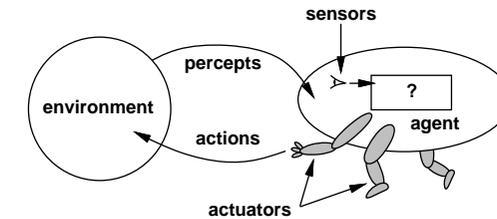
KI-Systeme erfüllen sehr **unterschiedliche** Aufgaben:

- ▶ Steuerung von Fertigungsanlagen
- ▶ Erkennen von Spam-Nachrichten
- ▶ Intralogistiksysteme in Lagerhäusern
- ▶ Einkaufsberatung im Internet
- ▶ Finden von Fehlern in Logik-Schaltkreisen
- ▶ ...

Wie fassen wir diese Vielfalt in einen **systematischen Rahmen**, der **Gemeinsamkeiten** und **Unterschiede** verdeutlicht?

Verbindende Metapher: **rationale Agenten** und ihre **Umgebungen**

Agenten



Agenten

- ▶ **Agentenfunktion** bildet Folgen von Wahrnehmungen (Perzepten) auf Aktionen ab:

$$f : \mathcal{P}^+ \rightarrow \mathcal{A}$$

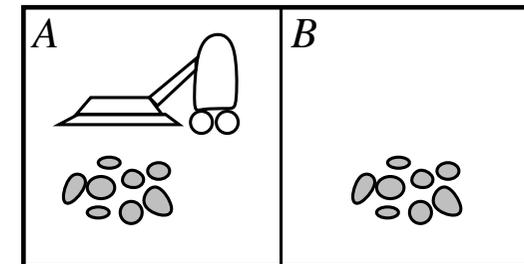
- ▶ **Agentenprogramm** läuft auf der physischen **Architektur** ab und berechnet f

Beispiele: Mensch, Roboter, Web-Crawler, Thermostat, OS-Scheduler, ...

Gestatten: ein Agent



Staubsaugerwelt



- ▶ **Wahrnehmungen:** Ort und Sauberkeit, z. B. $\langle a, \text{dirty} \rangle$
- ▶ **Aktionen:** left, right, suck, wait

Staubsauger-Agent

eine mögliche Agentenfunktion:

Wahrnehmungsfolge	Aktion
$\langle a, \text{clean} \rangle$	<i>right</i>
$\langle a, \text{dirty} \rangle$	<i>suck</i>
$\langle b, \text{clean} \rangle$	<i>left</i>
$\langle b, \text{dirty} \rangle$	<i>suck</i>
$\langle a, \text{clean} \rangle, \langle b, \text{clean} \rangle$	<i>left</i>
$\langle a, \text{clean} \rangle, \langle b, \text{dirty} \rangle$	<i>suck</i>
...	...

Reflexive Agenten

Bei **reflexiven** Agenten hängt die nächste Aktion nur vom **letzten** Perzept der Wahrnehmungsfolge ab:

- ▶ sehr einfaches Modell
- ▶ sehr eingeschränkt
- ▶ entspricht Mealy-Automaten (einer Art endlicher Automat) mit nur 1 Zustand
- ▶ praktische Beispiele?

Beispiel:

Reflexiver Staubsauger-Agent

```
def reflex-vacuum-agent(location, status):
    if status = dirty: return suck
    else if location = a: return right
    else if location = b: return left
```

Beurteilung von Agentenfunktionen

Was ist die **richtige** Agentenfunktion?

3.2 Rationalität

Rationalität

Rationales Verhalten

Verhalten von Agenten wird durch **Performance-Mass** bewertet (verwandte Begriffe: **Nutzen**, **Kosten**).

Perfekte Rationalität: wähle immer eine Aktion, die

- ▶ gegeben **verfügbare Informationen** (bisherige Perzepte)
- ▶ den **Erwartungswert** für **zukünftige Performance** maximiert

Frage: Ist der reflexive Staubsauger-Agent perfekt rational?

Hängt von Performance-Mass und Umgebung des Agenten ab!

- ▶ Haben Aktionen zuverlässig gewünschten Effekt?
- ▶ Kennen wir die Anfangssituation?
- ▶ Kann neuer Schmutz entstehen, während der Agent handelt?

Rationaler Staubsauger

Beispiel: Staubsauger-Agent

Performance-Mass:

- ▶ +100 Punkte pro gereinigtem Feld
- ▶ -10 Punkte pro *suck*-Aktion
- ▶ -1 Punkt pro *left/right*-Aktion

Umgebung:

- ▶ Aktionen und Wahrnehmungen zuverlässig
- ▶ Welt ändert sich nur durch Aktionen des Agenten
- ▶ alle Anfangssituationen gleich wahrscheinlich

Wie verhält sich ein perfekt rationaler Agent?

Rationalität: Diskussion

- ▶ **Perfekte Rationalität** \neq **Allwissen**
 - ▶ unvollständige Informationen (durch Perzepte) reduzieren erzielbaren Nutzen
- ▶ **Perfekte Rationalität** \neq **perfekte Vorhersage**
 - ▶ unsicheres Verhalten der Umgebung (z. B. stochastische Effekte von Aktionen) reduziert erzielbaren Nutzen
- ▶ **Perfekte Rationalität ist selten erreichbar**
 - ▶ beschränkte Rechenkapazität \rightsquigarrow **begrenzte Rationalität** (*bounded rationality*)

3.3 Zusammenfassung

Zusammenfassung (1)

verbindende Metapher für KI-Systeme: **rationale Agenten**

Agent interagiert mit **Umgebung**

- ▶ Sensoren erfassen **Wahrnehmungen** über Zustand der Umgebung
- ▶ Aktuatoren führen **Aktionen** aus, die Umgebung beeinflussen
- ▶ formal: **Agentenfunktion** bildet Perzeptfolge auf Aktionen ab
- ▶ **reflexiver** Agent: Agentenfunktion hängt nur von letzter Wahrnehmung ab

Zusammenfassung (2)

rationale Agenten:

- ▶ versuchen **Performance-Mass** zu maximieren
- ▶ **perfekte Rationalität**: Erreichen von maximaler Performance im Erwartungswert gegeben verfügbare Informationen
- ▶ bei „interessanten“ Problemen selten erreichbar:
begrenzte Rationalität